Thermistor at a Distance: Unobtrusive Measurement of Breathing

Jin Fei and Ioannis Pavlidis*, *Senior Member, IEEE*

Abstract—This paper unveils a thermal imaging methodology to recover the breathing waveform from the subject's nostrils. The resulting functionality is equivalent to that of a thermistor, but it is materialized in a contact-free manner. First, the nostril region is segmented and tracked over time through a network of cooperating probabilistic trackers. The mean thermal signal of the nostril region carries the breathing information. This information is extracted through wavelet analysis. The method has been tested on 20 healthy individuals. The breathing waveforms determined via the imaging computation were compared with the corresponding ones extracted from thermistors. The high degree of agreement between the two measurement methods confirms the validity of the proposed approach and opens the way for clinical applications. Furthermore, thermal imaging can be potentially used as an investigative tool to understand breathing physiology in ways not possible with contact sensors.

Index Terms—Breath, facial tracking, sleep studies, thermal imaging, wavelets.

Fig. 1. Subject undergoing sleep study. The discomfort from the nasal thermistor, the wires, and tapes is evident. The oxymoron is that subjects undergoing such monitoring are having trouble with sleep in the first place.

I. INTRODUCTION

ANALYSIS of breath waveforms plays an important role in the diagnosis and management of respiratory diseases like obstructive sleep apnea and asthma. In fact, breathing rate is one of the vital signs and hence indicative of the overall health status of a subject.

Human breathing consists of expiration and inspiration phases. The expired air has higher temperature than the inspired air due to heat exchange in the lungs and respiratory passageways [1], [2]. This thermic nature of breath around the nostril area creates an opportunity for a thermal measurement. It is the underlying principle of operation for the nasal thermistor. The thermistor probe has to be secured in front of the nostrils—a rather uncomfortable proposition.

Various other contact modalities have been developed to measure breath function that capitalize on different aspects of the breathing phenomenon. A very popular device is the respiratory belt transducer. It measures the breathing rhythm via pressure changes on the strap sensor fitted on the subject's chest or abdomen [3]. It is quite sensitive to motion.

There is a strong need for unobtrusive breathing measurement methods in certain clinical applications. Prime examples are sleep studies (see Fig. 1) and neonatal monitoring, where lengthy measurement of breath function with the minimum amount of discomfort is required. The first noncontact breathing rate measurement method was introduced by Greneker *et al.* [4] and is based on active sensing. It is called radar vital signs monitor (RVSM) and is able to measure the subject's heart beat and breathing rate at distances up to 30 ft. It senses the chest wall moving up and down during breathing by Doppler modulated radar. The RVSM measurements are sensitive to small body movement.

Thermal IR imaging is a passive contact-free modality. The sensing element itself can be viewed as a 2-D array of contact-free thermistors. In previous publications, we have demonstrated that thermal imaging can be used to measure various physiological variables, including blood flow [5], heart rate [6], [7], and breathing rate [8]. In fact, it is an ideal modality for sustained physiological monitoring [9].

We demonstrated for the first time, the feasibility of breath rate measurement through thermal imaging in [8]. Specifically, we proposed a statistical methodology that models breathing as a mixture of expiration and nonexpiration distributions. Every frame is classified as expiratory or nonexpiratory by comparing the incoming distributions with the existing distributions using the Jeffrey's divergence measure. Due to this frame labeling, we are able to compute the breathing rate. We retrofitted in [10] the thermal imaging sensor with an optical bandpass filter in the CO_2 absorption zone to improve the SNR. Then we proceeded to compute the breathing rate using Fourier analysis. In these early efforts, the measurement region was selected manually and

there was inadequate compensation for motion. The subjects were instructed to stay as still as possible in the course of the experiments.

In this paper, we introduce a new and improved breathing measurement method based on automatic tracking/localization of the nasal region and wavelet analysis. No bandpass filtering is necessary. Not only the breathing rate, but also the full breathing waveform is recovered this time. Specifically, in Section II, we describe the tracking and localization algorithms as well as the wavelet-based signal extraction method. We discuss the experimental setup in Section III-A, sensitivity analysis in Section III-B–C, and the experimental results in Section III-D. Section IV concludes the paper.

II. METHODOLOGY

To measure the breathing function in thermal video, we need to track the motion of subject, localize the measurement region, and analyze the extracted signal. We address each of these issues in detail in the following sections.

A. Tracking

We chose the coalitional tracking algorithm [11], [12] to track facial tissue during breath measurements. It optimizes multi-tracker interaction via game theory. The coalitional tracking method was developed to address the conflicting goals of generality and accuracy that arise in the context of thermophysiological measurements on the face. Thermal imaging is a functional imaging that depicts an evolving physiological process. The dynamic nature of thermal imaging poses a modeling challenge to track. Particle filter trackers [13], [14] overcome this challenge because they are general and adapt well to changes. The accuracy of these trackers peaks when the real estate they cover is neither too large nor too small. By optimizing the behavior of a spatially distributed cluster of particle filter trackers (coalition), one gains in accuracy without sacrificing adaptability.

The characteristic function of a tracking coalition C_k at time t is

$$\Pi_{C_k}^{(t)} = \omega_\alpha \times \alpha_{C_k}^{(t)} + \omega_\beta \times \beta_{C_k}^{(t)} + \omega_\gamma \times \gamma_{C_k}^{(t)} + \omega_\delta \times \delta_{C_k}^{(t)} \quad (1)$$

$$\omega_\alpha + \omega_\beta + \omega_\gamma + \omega_\delta = 1 \quad (2)$$

$$\omega_\alpha, \omega_\beta, \omega_\gamma, \omega_\delta \in [0, 1] \quad (3)$$

where $\alpha_{C_k}^{(t)}, \beta_{C_k}^{(t)}, \gamma_{C_k}^{(t)},$ and $\delta_{C_k}^{(t)}$ are the four characteristic scores, and $\omega_\alpha, \omega_\beta, \omega_\gamma,$ and ω_δ are the corresponding weights.

1) $\alpha^{(t)}$ is the dynamic template match score, which rewards trackers that maintain consistent imaging content.
2) $\beta^{(t)}$ is the geometric alignment score that favors coalitions whose members have geometric alignment analogous to the original ($t = 0$) configuration.
3) $\gamma^{(t)}$ is the interframe projection agreement score; it is a continuity constraint to improve robustness by penalizing abrupt (and improbable) changes of the projected state of the target between successive frames.
4) $\delta^{(t)}$ is the interframe membership retention score. It is also a continuity constraint to reflect the tendency of the
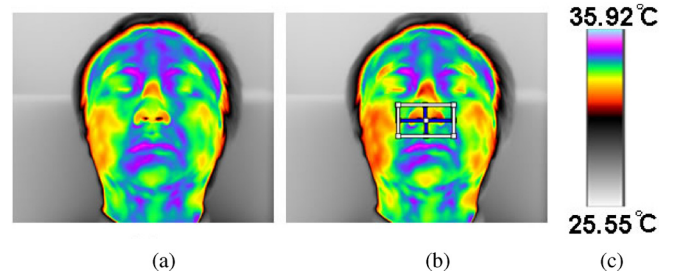

Fig. 2. (a) Thermal snapshot of a subject's face. (b) Initialization of coalitional tracker (TROI). (c) Thermal color map.

winning coalition from the previous time step to retain its members.

We have $\binom{N}{k}$ different coalitions of size k ($k \in 1, 2, \ldots, N$), out of which we select the one with the highest score. We define this as the winning coalition C^t at time t

$$C^t = \underset{C_k}{\operatorname{argmax}} \, \Pi_{C_k}^{(t)}. \quad (4)$$

We use a coalition grid composed of four-particle filter trackers. The grid outline is drawn interactively by a click and drag operation on the first frame. It encompasses comfortably the nostril region without any stricter specification. The grid outline constitutes the tracking region of interest (TROI) (see white rectangle in Fig. 2). The localization algorithm (see Section II-B) operates within TROI and determines the nostril region where the measurement is performed more rigorously.

B. Localization

The source of thermal signal, and therefore, the measurement ROI (MROI) for breathing is the nostrils. This region features both spatial and temporal variances. First, the shape of the nostril region is different for different individuals. Thermal imaging is a functional modality that records the changing physiology. In the case of breathing, thermal imagery registers the temperature fluctuation between the inspiration and expiration phases. When the subject inspires, there is suction of cooler air from the environment, which convects away heat from the nostril tissue at a high rate. As a result, the nostril area radiates toward the camera at a lower power. When the subject expires, there is forceful ejection of warm air toward the environment. Since the temperature difference between the expired air and the nostril tissue is minimal this time, convection is weak and the thermal signature of the nostril dominates in the imagery. As a result, the nostril area radiates toward the camera at a higher power. The phenomenon can be clearly visualized in Fig. 3. It is interesting to observe that although the nostrils become hotter during expiration, the outer cartilage at the bottom of the nose becomes colder [see Fig. 3(c)]. This is because the progressively cooled expiratory air is funneled upward, abducting away heat from the nose's outer sides. The intermediate apneic phase between inspiration and expiration [see Fig. 3(b)] allows the true tissue radiation to show; thus, both the nostril and outer cartilage areas appear hot.
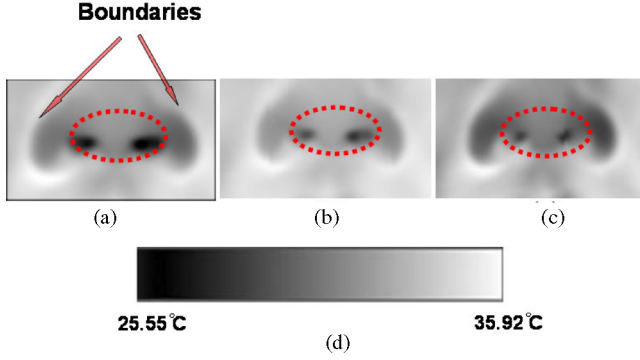
Fig. 3.   Temporal variance of nostril region in thermal imagery during breathing. (a) Inspiration phase. (b) Transition phase. (c) Expiration phase. (d) Thermal color map.
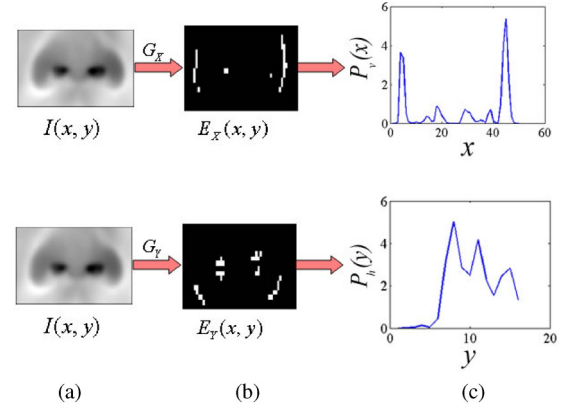

Fig. 4.   (Top) Determination of left and right outer nostril edges. (Bottom) Determination of base edge. (a) TROI images. (b) Edge images. (c) Integral projections.
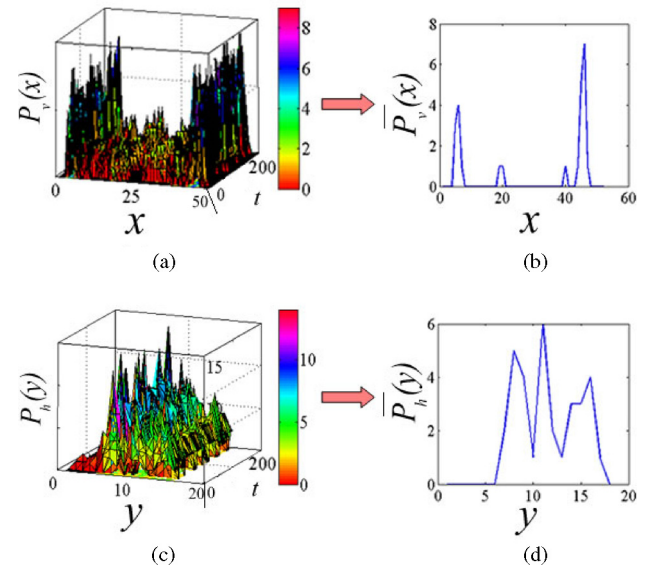

Fig. 5.   Spatiotemporal evolution of integral nostril projections for a typical subject. (a) Vertical projections along time. (b) Mean vertical integral projection. (c) Horizontal projections along time. (d) Mean horizontal integral projection.

Since the shape of nostrils varies temporally due to the varying thermal signature of inspiration and expiration, segmentation is challenging. Fig. 3 shows how within TROI, the nostrils are separated from the rest of the facial tissue due to colder boundaries formed by the cartilage. The contrast varies over the breathing cycle, but it never goes away altogether. This feature can help to localize MROI. The contrast at the boundary of the nose is quite strong not only in thermal imagery, but also in visual imagery, for different reasons. The nose is a distinct 3-D feature in an otherwise 2-D facial surface and forms strong edges at the seams. A lot of research has focused on feature extraction for face recognition [15]–[20]. Due to similar nose boundary properties, we can leverage some of the works performed in the visual spectrum for thermal imaging purposes. Specifically, Brunelli and Poggio [16] showed that the horizontal gradients are useful in detecting the left and right visual boundaries of nose, whereas vertical gradients can detect the nose base. Kotropoulos and Pitas [16] demonstrated that the vertical and horizontal projection profiles of human nose are obtained by summing-up visual pixel intensities row-wise and column-wise, respectively. We have used elements of these approaches transplanted in the thermal IR domain.

Let $I(x, y)$ be the original thermal image, and $E_X(x, y)$ and $E_Y(x, y)$ be the edge images after applying the following vertical (G_X) and horizontal (G_Y) Sobel edge templates:

$$G_X = \begin{pmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{pmatrix} \qquad G_Y = \begin{pmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}.$$

We perform integral projections on the edge images to extract the facial features of interest, i.e., the nostrils' outer edges. The vertical integral projection is defined as

$$P_v(x) = \sum_{y=y_1}^{y_m} E_X(x, y). \tag{5}$$

Similarly, the horizontal integral projection is defined as

$$P_h(y) = \sum_{x=x_1}^{x_n} E_Y(x, y). \tag{6}$$

The left- and rightmost peaks of $P_v(x)$ yield the left and right outer nostril edges correspondingly. The maximum of $P_h(y)$ yields the base edge (see Fig. 4).

The MROI selection varies from frame to frame. Fig. 5(a) and (c) shows the vertical and horizontal nostril projections along the time line for a typical subject. Some projections are weak and the locations also vary. We use a time window of 200 frames (4–5 s) to compute the mean vertical and horizontal projections in Fig. 5(b) and (d), respectively. This time window is a representative of the full spatiotemporal evolution as it covers both expiration and inspiration phases.

Based on anthropometric knowledge, we estimate the nostril's height H as one third of its base edge segment W [21], which is delineated by the left and right outer nostril edges. Thus, we construct the MROI $W \times H$ (see Fig. 6). We compute the mean temperature within MROI in every frame. This produces a quasi-periodic thermal signal along the time line, which is indicative of the breathing function.
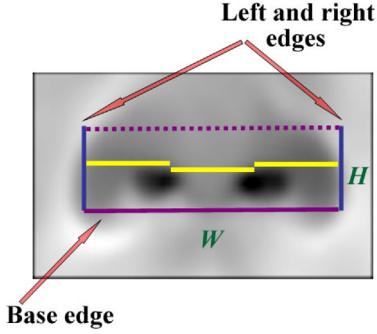
Fig. 6. MROI based on mean nose edges and anthropometric estimates.

C. Wavelet Analysis

We perform identical wavelet analysis on the thermal imaging and thermistor nasal signals to recover the breathing waveform and rate. Thus, it is sufficient to present the imaging nasal signal analysis only. Wavelets are the appropriate analysis tool as breathing is a nonstationary process.

The thermal video sampling rate fluctuates around 55 frames per second (fps). This is due to the use of a standard PC instead of a hard real-time platform for image acquisition. A constant sampling rate is necessary for optimal results in wavelet decomposition. We choose $\delta = 10$ fps as the resampling rate of the thermal signal.

We normalize and perform wavelet analysis on sliding segments (windows) of the resampled thermal signal. As the sliding window travels along the evolving time line of the resampled and normalized signal, we compute a series of breathing waveforms and rates. This renders the breath computation near real time. The details of each algorithmic step are as follows.

1) Normalization: We define as $S_g(t)$, $t \in \{0, \ldots, N_g\}$, the resampled breathing signal on the segment g. We normalize the signal amplitude as follows:

$$S'_g(t) = \frac{S_g(t) - \mu_g}{\sigma_g} \qquad (7)$$

where μ_g and σ_g are the mean and standard deviation of $S_g(t)$, respectively. The normalization transforms signal $S_g(t)$ to $S'_g(t)$ with mean $\mu'_g = 0$ and standard deviation $\sigma'_g = 1$.

2) Wavelet Transform: We perform continuous wavelet transformation (CWT) [22] on the resampled and normalized thermal signal segment

$$\Psi^{\Psi}_{S'_g}(\tau, s) = \frac{1}{\sqrt{|s|}} \int S'_g(t) \psi\left(\frac{t-\tau}{s}\right) dt \qquad (8)$$

where ψ is the "mother wavelet," τ represents the translation parameter, while s denotes the scale at which the signal segment is examined. We use the Mexican Hat (MH) [22] as the mother wavelet.

3) Breathing Waveform: CWT allows analysis at all scales, thus facilitating the extraction of the signal component of interest (i.e., breathing). We assume that the breathing component exists at a scale $s_{\max}$ corresponding to a local maximum of the

Fig. 7. Experimental setup: subject, imaging, and ground-truthing clinical equipment.

wavelet energy coefficients $WT_i(t)$

$$s_{\max} = \underset{i}{\operatorname{argmax}} \left\{ \sum |WT_i(t)|^2 \right\}. \qquad (9)$$

Lower scales than $s_{\max}$ are likely to contain noise, while higher scales contain metabolic contributions. The comparative experimental results in Section III-D verify the correctness of this hypothesis; thus, it can be safely used to compute the breathing waveform in thermal imagery of the nostrils.

4) Breathing Rate: Given a mother wavelet, the frequency that maximizes its transform is the center frequency F_c. This frequency is known for the MH mother wavelet used to represent the generic nostril signal. The center frequency F_c adjusted for downsampling $\delta = 10$ fps is the breathing frequency BR dilated (scaled) by the factor $s_{\max}$

$$\text{BR} \cdot s_{\max} = F_c \delta. \qquad (10)$$

One can solve this equation to find the breathing rate BR.

III. EXPERIMENTS

A. Experimental Setup

The centerpiece of the imaging system we used in our experiments is a forward looking IR (FLIR) SC6000 midwave IR (MWIR) camera with an indium antimonite (InSb) detector operating in the range 3–5 μm [23]. The camera has a focal plane array (FPA) with maximum resolution of 640×512 pixels. The sensitivity is $0.025\,^{\circ}$C. The camera is outfitted with an MWIR 50-mm lens $f/2.3$, $Si : Ge$, bayonet mount from FLIR Systems [23].

The ground truth system is composed of a PowerLab 8/30 data acquisition system and a thermistor from ADInstruments [3]. An electronic trigger synchronizes the imaging and ground truth systems.

The experiments took place in a climate-controlled room according to an approved protocol by the Institutional Review Board, University of Houston. Subjects were located 6 ft away from the imaging system and offered a frontal view when sitting in a comfortable chair (see Fig. 7). The subjects were also
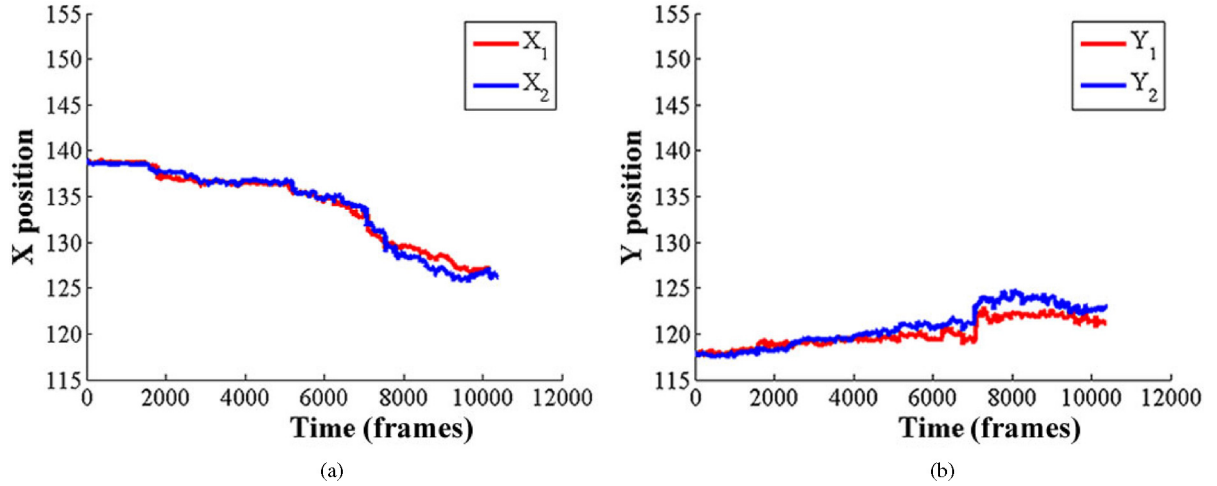
Fig. 8. Evolution of tracker's centroid in two runs for a typical thermal clip.

fitted with the nasal thermistor to ground truth the imaging measurements. The MWIR camera was calibrated with a two-point calibration at 28 °C and 38 °C, which are the end points of a typical temperature distribution on a human face. We recorded 3 min thermal clips (and corresponding thermistor signals) for 20 healthy subjects of both sexes.

The subjects were left free to move in their chairs. During the recording, some of them appeared more agitated than others, thus providing good testing data for the tracker. Almost all subjects exhibited the normal breathing rates of resting adults (i.e., $12-20$ cycles per minute [24]), as validation of their respiratory health.

B. Sensitivity to Tracking

The coalitional tracker is composed of particle filter trackers that have stochastic components. This means that the tracker's behavior is not absolutely deterministic. If the same tracker is run twice with exactly the same initialization conditions, it will produce slightly different results. Provided that repeatability stays within bounds, this is an acceptable compromise for the tracker to overcome local extrema and discretize effectively the search space.

To quantify the repeatability of the coalitional tracker, we ran it twice on all 20 thermal clips of the dataset. For each clip, the initialization conditions between runs remained the same. Fig. 8(a) and (b) shows a typical evolution of the tracker's centroid coordinates X_i and Y_i, respectively, where i denotes the run number.

We computed the Euclidean distances $\text{Dist}_{X_t} = \|X_{1t} - X_{2t}\|$ and $\text{Dist}_{Y_t} = \|Y_{1t} - Y_{2t}\|$ of the tracker's centroid along the time line $t \in \{0, \ldots, N\}$. Table I shows the first- and second-order moments for the centroid drift between runs. On average, the drift is negligible as it remains within 1 pixel in either dimension. Only in the case of $S19$, where the tracking performance is not effective, the drift is somewhat larger.

TABLE I
REPEATABILITY STATISTICS FOR TRACKER CENTROID

Subject	μ_{Dist_X}	σ_{Dist_X}	μ_{Dist_Y}	σ_{Dist_X}
S01	0.48	0.44	0.93	0.66
S02	0.09	0.07	0.26	0.15
S03	0.18	0.15	0.15	0.10
S04	0.48	0.43	0.28	0.19
S05	0.22	0.16	0.32	0.25
S06	0.07	0.06	0.17	0.12
S07	0.16	0.10	0.49	0.36
S08	0.74	0.36	0.23	0.15
S09	0.44	0.39	0.31	0.26
S10	0.26	0.18	0.22	0.19
S11	0.30	0.24	0.52	0.41
S12	0.16	0.10	0.29	0.19
S13	0.35	0.30	1.00	0.78
S14	0.47	0.28	0.54	0.38
S15	1.53	1.66	0.70	0.51
S16	0.27	0.22	0.43	0.35
S17	0.36	0.25	0.67	0.43
S18	0.44	0.34	0.62	0.47
S19	2.73	3.15	0.79	0.82
S20	0.23	0.18	0.74	0.61
Average	0.50	0.45	0.48	0.37

C. Sensitivity to MROI Selection

In Section II-B, we described an algorithm for automatic MROI localization. The estimated MROI is meant to cover the outer extent of the nostril region. Strictly speaking, the source of the signal is the nostril orifices. The MROI covers nearby cartilage in addition to the orifices. This simplifies segmentation and provides some room for tracking error. The question then becomes how sensitive the breathing measurement is to liberal MROI selection.

To quantify sensitivity, we chose a random clip from the dataset, and generated slightly larger and smaller MROIs from a baseline one. In Fig. 9, the baseline MROI is shown in solid yellow, and it is the minimal rectangle covering tightly the two nostrils. The dashed pink and blue MROIs are offset inward by 1 and 2 pixels, respectively. Similarly, we have four MROIs offset
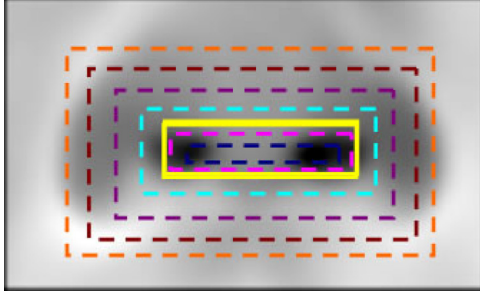
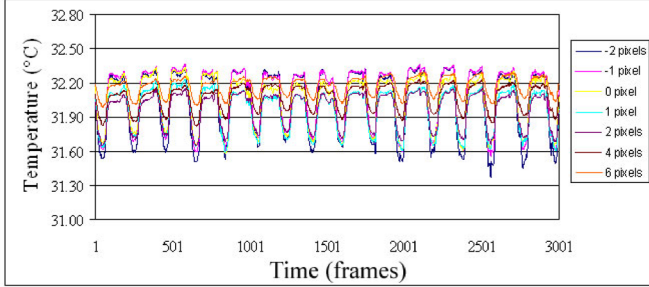Fig. 9. Various size MROIs centered on the nostrils.

Fig. 10. Thermal signals corresponding to the MROIs depicted in Fig. 9.

Fig. 11. MROI localization for different TROIs.

outward by 1, 2, 4, and 6 pixels correspondingly. The latter one is the MROI generated automatically by the localization algorithm described in Section II-B.

Fig. 10 shows the extracted mean thermal signals from the seven MROIs of Fig. 9. The waveform colors correspond to the MROI colors. We observe that the signals have identical periodicity and very similar amplitude. Therefore, slightly larger or smaller MROIs from the baseline can preserve the source breath signals as long as they are centered on the nostrils.

In addition, we examined the robustness of the MROI localization algorithm by selecting various sizes for TROI. Naturally, TROI differs from trial to trial and from user to user. It is meant to serve as gross initialization in the absence of a face segmenter. Despite the inherent variance in the TROI selection, the MROI algorithm should always provide precisely the same region on a particular clip. For every subject in the dataset, we selected six different TROIs and ran a corresponding number of times the MROI algorithm. We also manually delineated the baseline MROI that encompasses tightly the nostril orifices. Fig. 11 shows visually the results for subject S_{15}.

A measure of consistency is for the automatic MROI to include entirely the baseline MROI to prevent any loss of signal information. We define as coverage C_n

$$C_n = \frac{R(\text{Automatic MROI} \cap \text{Baseline MROI})}{R(\text{Baseline MROI})} \quad (11)$$

where $R(.)$ denotes the area of a region. Table II shows the computed C_n for all six TROI trials in the entire dataset. On average, the coverage is close to 95%. The MROI algorithm performs better when the user chooses smaller TROI (e.g., TROI 1) and deteriorates somewhat when he/she chooses larger TROI (e.g., TROI 6).
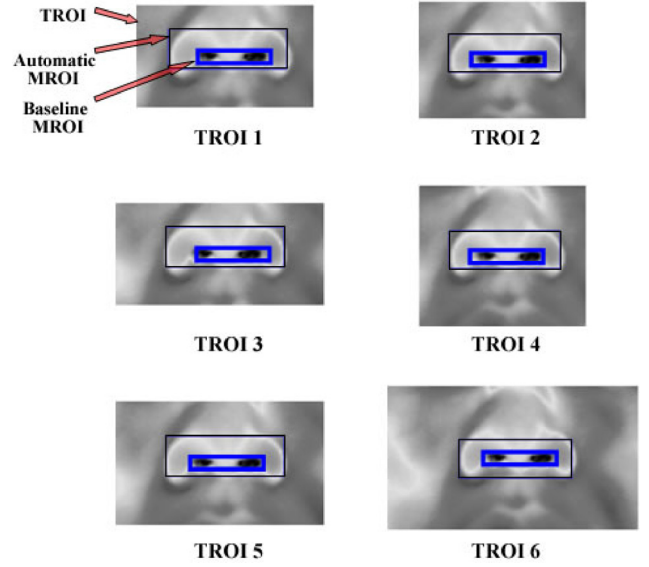
TABLE II
COVERAGE C_n (IN PERCENT) OF TROI 1–6

Subject	TROI 1	TROI 2	TROI 3	TROI 4	TROI 5	TROI 6
S01	100	100	100	100	100	100
S02	100	100	100	100	100	100
S03	100	100	100	100	100	100
S04	95	95	95	95	95	95
S05	100	100	100	100	100	100
S06	100	100	100	91	100	100
S07	100	100	100	100	100	100
S08	91	91	100	91	56	100
S09	100	100	100	100	100	100
S10	100	100	72	72	28	12
S11	100	100	100	100	100	100
S12	100	100	100	100	72	72
S13	100	100	100	97	100	51
S14	100	78	100	78	100	100
S15	100	100	100	100	100	100
S16	99	99	100	100	99	100
S17	100	100	100	100	100	100
S18	100	100	100	100	100	100
S19	100	100	100	100	100	100
S20	100	100	100	100	100	59
Average	99	98	98	96	92	89

D. Experimental Results

Based on the methodology described in Section II, we determine the breathing waveforms from the thermal imaging and thermistor signals by computing the wavelet energy curves in all scales, and selecting the local maximum in the small-scale region. The small-scale region corresponds to relatively high-frequency phenomena embedded in the thermal signals, and breathing is one of them. Actually, it is expected to be the strongest phenomenon present due to the measurement locale. Fig. 12 shows the imaging and thermistor energy curves for the corresponding signals of all subjects included in the test population. One can observe the simultaneous peaks of the two curves,
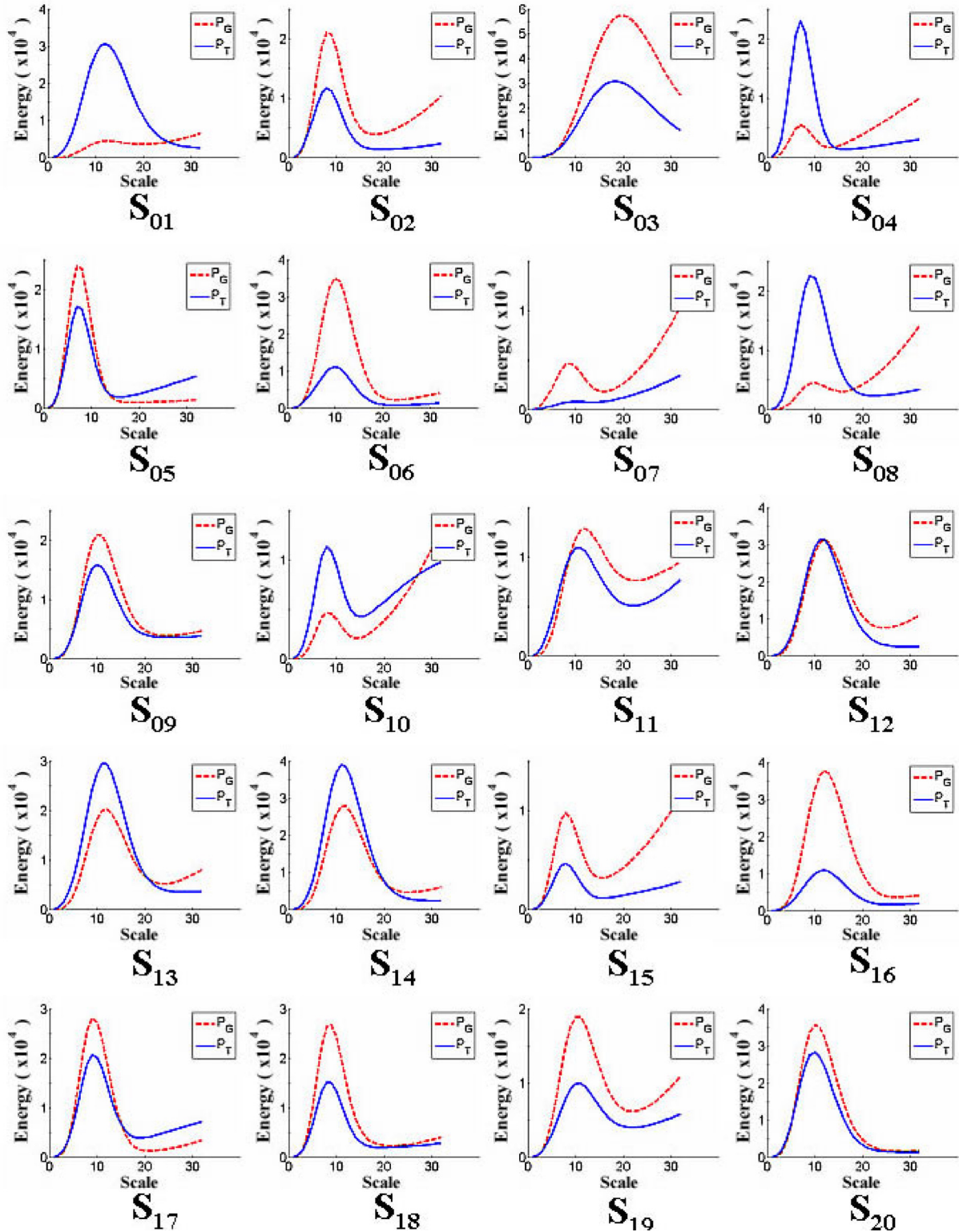
Fig. 12.   Wavelet energy curves of the thermal imaging P_T and thermistor P_G signals for all subjects. Only small scales [1]–[32] are shown.

which indicate agreement between the two modalities as to the scale (and thus, the frequency) of the breathing phenomenon.

Fig. 13(a) shows the thermal imaging and ground truth wavelets corresponding to the breathing scale $s_{max} = 10$ for subject S_{20}. Fig. 13(b) shows the cross correlation (XR) be-

tween the two waveforms as a function of time lag. The maximum XR is 97.48%, where the thermally imaged breathing waveform $S_{s_{max}}^{T}(t)$ is -4 lags, i.e., 0.4 s behind the ground truthed one, i.e., $S_{s_{max}}^{G}(t)$. Table III shows that, on average, the maximum XR in the dataset is 92.75%. Overall, the breathing
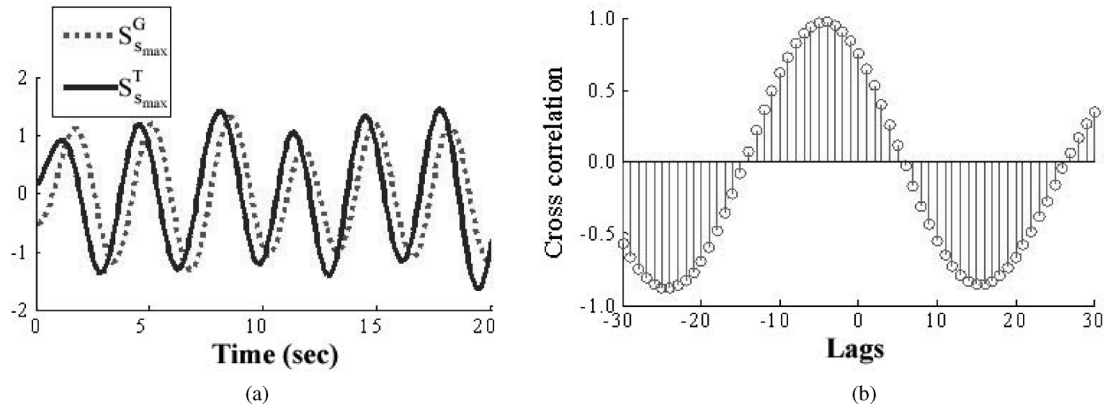
Fig. 13.   (a) Breathing wavelets of imaged and ground truth signals. (b) Cross correlation.

TABLE III
XR OF BREATHING WAVEFORMS

Subjects	s_{max}^T	s_{max}^G	$Lag(sec)$	$XR(\%)$
S01	12	13	-1.20	82.15
S02	8	8	-0.50	96.74
S03	18	20	-0.80	93.72
S04	7	7	-0.40	96.12
S05	7	7	-0.30	97.23
S06	10	10	-0.50	97.73
S07	10	9	0.40	77.39
S08	9	10	-0.70	90.40
S09	10	10	-0.40	93.23
S10	8	8	-0.60	89.18
S11	11	12	-0.60	88.23
S12	12	12	-0.60	96.11
S13	12	12	-0.60	96.08
S14	11	12	-0.60	96.25
S15	8	8	-0.50	92.26
S16	12	12	-0.50	96.40
S17	9	9	-0.50	95.30
S18	8	9	-0.40	95.75
S19	11	11	0.00	77.64
S20	10	10	-0.40	97.48
Mean	10	10	-0.48	92.27

TABLE IV
BREATHING RATE RESULTS

Subjects	$\overline{BR_G}$	$\overline{BR_T}$	$CAND(\%)$	$CuSum(\%)$
S01	12.83	12.81	99.81	13.27
S02	18.49	18.34	99.19	6.80
S03	8.22	8.33	98.68	10.37
S04	21.52	21.23	98.64	7.45
S05	20.76	21.17	98.03	6.80
S06	14.89	14.88	99.97	8.26
S07	16.56	16.68	99.26	6.88
S08	16.08	16.22	99.09	13.10
S09	15.20	14.80	97.39	11.99
S10	18.70	18.69	99.95	7.24
S11	13.55	13.22	97.54	19.54
S12	13.30	13.10	98.49	11.27
S13	13.31	13.06	98.07	13.65
S14	13.67	13.52	98.89	9.04
S15	21.65	21.88	98.92	9.46
S16	12.75	11.63	91.25	13.40
S17	16.41	16.23	98.94	6.33
S18	17.72	17.55	99.06	10.06
S19	14.50	13.99	96.48	11.25
S20	15.59	15.25	97.83	12.44
Mean	15.79	15.63	98.27	10.42

waveform extracted from the thermal imaging data lags the breathing waveform extracted from the thermistor by -0.48 s. This small delay is due to the imaging computation and insignificant. As the exceptionally effective scale agreement and XR numbers demonstrate (second, third, and fifth columns in Table III), the extracted breathing waveforms from the two modalities are almost identical. This is a firm confirmation that the thermal imaging method can function as a virtual thermistor.

Table IV shows the detailed experimental results regarding the rates of the breathing signals. The imaging measurements (BR_T) are juxtaposed with the corresponding ground truth ones (BR_G) obtained through the thermistor.

The comparison between the two modalities is based on two measures, CAND and CuSum. CAND stands for complement of the absolute normalized difference given by

$$CAND = 1 - \frac{|\overline{BR_T} - \overline{BR_G}|}{\overline{BR_G}} \qquad (12)$$

which is the absolute difference between the mean thermal imaging and ground truth measurements normalized against the mean ground truth and subtracted from unity. This gives a weighted indication of how close the mean thermal imaging measurement ($\overline{BR_T}$) is to the mean ground truth measurement ($\overline{BR_G}$) in each case. The mean $\overline{CAND}$ for the experiment is 98.27%.

CuSum stands for cumulative sum, and is an indication of the instantaneous error in the imaging measurement with respect to the ground truth measurement

$$CuSum = \frac{1}{T}\sum_{t=1}^{T}\frac{|BR_T(t) - BR_G(t)|}{BR_G(t)} \times 100\%. \qquad (13)$$

The mean $\overline{CuSum}$ error for the experiment is 10.42%. On average (see $\overline{CAND}$), the two modalities give almost identical measurements. There is more discrepancy if one examines the measurements on a beat-by-beat basis (see $\overline{CuSum}$), but these tend to cancel out on the whole.
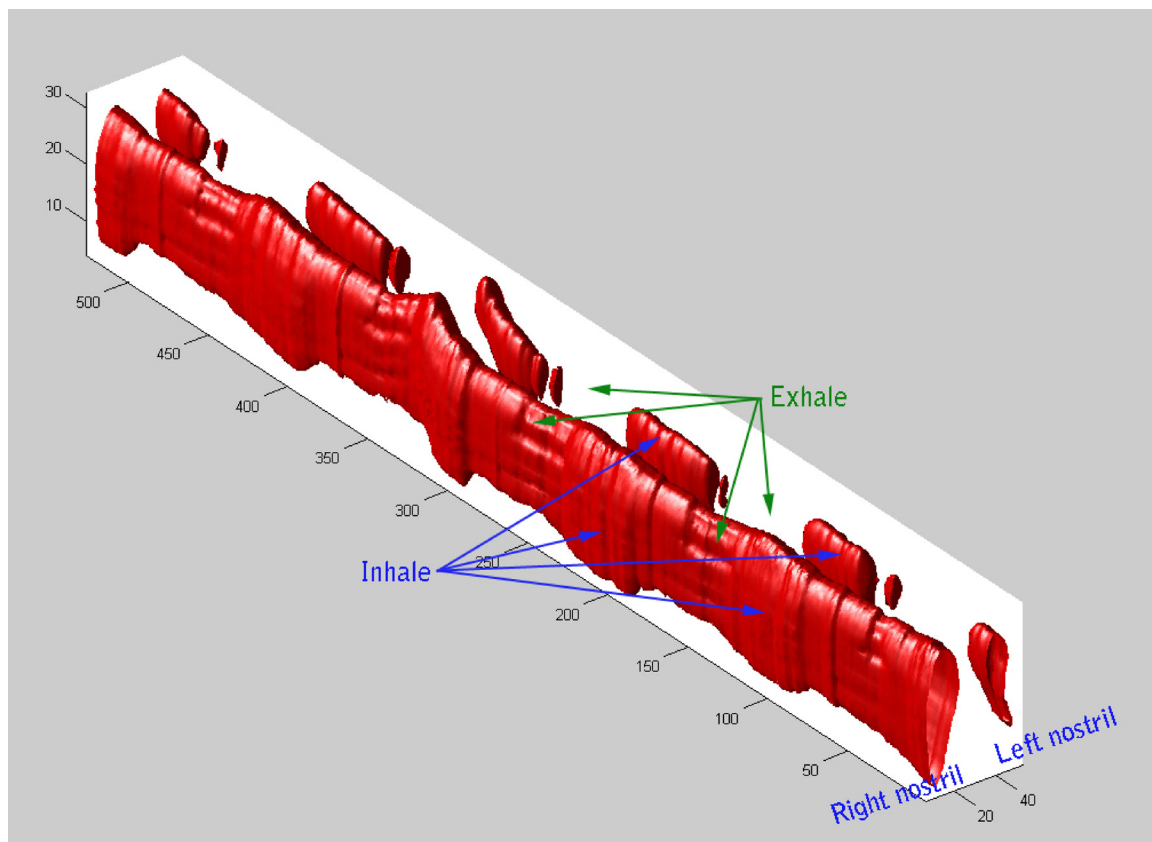
Fig. 14. Visualization of 2-D thermal isosurfaces in the nostrils over time for a subject. This is the native information the thermal imaging sensor produces and one can intuitively understand the wealth of information hidden therein. However, mathematical handling of this information becomes more involved. In the current study, a mean value was obtained at each point in time to form the 1-D signal that emulates a thermistor operation.

IV. DISCUSSION

Breathing is one of the vital signs. It is used as an indicator of overall health status and also in diagnosis of chronic or acute diseases, like obstructive sleep apnea or heart attack. In certain cases (e.g., sleep studies or neonatal care), extensive monitoring of breathing function is necessary in as unobtrusive manner as possible. In this paper, we have described a new imaging method that fulfills these demanding specifications. It is a leap forward with respect to previous methods we reported on this matter. The present method automatically localizes the nostril region and cancels the effect of head motion. Also, a novel departure from previous methods that we reported is the use of wavelet analysis for the determination of not only the breathing rate, but also the full breathing waveform. This is a better approach than Fourier analysis as breathing is a nonstationary process. Based on the largest scale experiment we conducted yet, the thermal imaging measurement appears to be in par with the existing clinical standard. Now that the feasibility of breath measurements with passive imaging means has been firmly established, future efforts should be concentrated on pathophysiological experiments.

One could ponder various questions that arise from the prospect of applying the method in clinical practice, particularly in sleep studies. At closure, we briefly discuss a few of these questions in an effort to render further help in future clinical investigations.

1) *Cost:* The thermal imaging system has high initial cost. However, it has negligible maintenance cost once it is acquired. This is in complete antithesis to a classic polysomnographic system, which has lower initial cost, but higher operational cost. The latter stems from the sterilization and replacement of consumables. Therefore, over the long run, the cost-effectiveness of the thermal imaging solution is viable. In fact, lower grade (and thus cheaper) thermal imaging systems can be probably used with equal effectiveness to the one used in this study. The reason is that the thermic effect of breathing appears to be in the tenths, not hundredths, of degrees of Celsius, and hence capable to afford lower sensitivity. Furthermore, the spatial extent of the phenomenon spreads over dozens of pixels, and hence capable to afford lower spatial resolution.

2) *Single channel:* Air flow is one of many channels used in polysomnography. However, it is the most important. There is a possibility for some additional channels to be proved measurable via thermal imaging through further research. For example, eye movements could be indirectly quantified through the metabolic heat signature of the ocular muscle. Some relevant work for a different application has been done in [25]. And pulsation has proved quantifiable through thermal imaging of superficial vasculature [7]. In general, since the entire face is imaged, there

is potential for virtual measurements of multiple physiological variables at appropriate locales.

3) *Operational envelope:* During sleep, the subject can turn and assume either a prone or lateral posture. In the case of classic polysomnography, the wires are likely to be severed in such turns and the sensors to be detached. The sleep study operator, who monitors the signals in another room, has to go to the patient's bedside and reattach the sensors or content with losing the signal for a period of the study. In the case of thermal imaging, if the subject turns on the side, the operator would simply need to relocate the virtual measurement area on the imagery using the mouse. This is a significant advantage. If the subject turns and assumes a prone posture, the operator would have to do what usually is done in polysomnographic studies—go to the patient's room and help him/her turning back (without reattaching wires this time) or content with losing the signal for a period of the study.

In general, there is considerable operational flexibility with the imaging method that does not exist in polysomnography. With imaging, when things go wrong, automation may break down, but the operator can click away from the problem, right at his/her monitor. With polysomnography, when things go wrong, the operator always have to rush to the patient's bedside and take physical corrective action. For example, if the nostril segmentation fails, because the subject turns on the side, or it is irrelevant, because the subject starts breathing from the mouth, the operator can fix the problem by simply repositioning the TROI and MROI areas with the mouse. Future algorithmic enhancements may improve robustness so that automation remains uninterrupted even when turns occur.

Beyond the potential clinical applications, this technology can be useful as an investigative tool to understand respiratory physiology deeper and improve current practices. This is especially true if one does not reduce the inherently multidimensional imaging signal to 1-D, as the method presented in this paper did. Fig. 14 shows the evolution of 2-D thermal isosurfaces in the nostrils of a subject over time. Breathing from the left nostril of the subject appears to be impaired. However, this subtle abnormality was masked in the thermistor data, as the thermistor was averaging from the two nostrils due to its in-between placement. This is a rather standard practice in sleep clinics, which should probably stop. It is far better to place one thermistor in each nostril, if one wants to catch subtle breathing abnormalities that are not shared by both nostrils. In other words, thermal imaging technology can be used to better understand what current simpler sensing counterparts are missing; this may lead to the adoption of optimization strategies, like the one presented in this example. One major contribution of this paper is that it opens the way for investigations like this to take place in the future.

ACKNOWLEDGMENT

REFERENCES

[1] Y. Zhang, "Quantitative measurement of radiation properties for opaque and semi-transparent greybodies," *Infrared Phys.*, vol. 30, no. 2, pp. 149–153, 1990.

[2] J. Lindemann, R. Leiacker, G. Rettinger, and T. Keck, "Nasal mucosal temperature during respiration," *Clin. Otolaryngol.*, vol. 27, no. 3, pp. 135–139, Jun. 2002.

[3] (2007). ADInstruments, 2205 Executive Circle, Colorado Springs, CO [Online]. Available: http://www.adinstruments.com

[4] E. Greneker, "Radar sensing of heartbeat and respiration at a distance with security applications," in *Proc. SPIE, Radar Sens. Technol. II*, R. Trebits and J. Kurtz, Eds., Jun. 1997, vol. 3066, pp. 22–27.

[5] M. Garbey, A. Merla, and I. Pavlidis, "Estimation of blood flow speed and vessel location from thermal video," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Washington, DC, Jun. 27–Jul. 2, 2004, vol. 1, pp. 356–363.

[6] N. Sun, I. Pavlidis, M. Garbey, and J. Fei, "Harvesting the thermal cardiac pulse signal," in *International Society and Conference Series on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (ser. Lecture Notes in Computer Science 4191). Copenhagen, Denmark: Springer-Verlag, 2006, pp. 569–576.

[7] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1418–1426, Aug. 2007.

[8] R. Murthy, I. Pavlidis, and P. Tsiamyrtzis, "Touchless monitoring of breathing function," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Fransisco, CA, Sep. 2004, vol. 2, pp. 1196–1199

[9] I. Pavlidis, "Continuous physiological monitoring," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Cancun, Mexico, Sep. 17–21, 2003, vol. 2, pp. 1084–1087.

[10] J. Fei, Z. Zhu, and I. Pavlidis, "Imaging breathing rate in the CO_2 absorption band," in *Proc. 27th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Shanghai, China, Sep. 1–4, 2005, pp. 700–705.

[11] J. Dowdall, I. Pavlidis, and P. Tsiamyrtzis, "Coalitional tracking," *Comput. Vis. Image Understanding*, vol. 106, no. 2/3, pp. 205–219, May/Jun. 2007.

[12] J. Dowdall, I. Pavlidis, and P. Tsiamyrtzis, "Coalitional tracking in facial infrared imaging and beyond," in *Proc. 2006 IEEE Conf. Comput. Vis. Pattern Recognit. Workshop (CVPRW)*, New York, Jun. 17–22, 2006, pp. 134–141.

[13] M. Isard and A. Blake, "Condensation—Conditional density propagation for visual tracking," *Int. J. Comput. Vis.*, vol. 29, no. 1, pp. 5–28, Aug. 1998.

[14] M. Isard and A. Blake, "Condensation: Unifying low-level and high-level tracking in a stochastic framework," in *Proc. 5th Eur. Conf. Comput. Vision (ECCV)* (ser. Lecture Notes in Computer Science 1406). Freiburg, Germany: Springer-Verlag, Jun. 2–6, 1998, pp. 893–908.

[15] W. Zhao and R.Chellappa, A.Rosenfeld, J. Phillips, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 25, no. 4, pp. 399–458, Dec. 2003.

[16] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 10, pp. 1042–1052, Oct. 1993.

[17] B. Low and M. Ibrahim, "A fast and accurate algorithm for facial feature segmentation," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, Lausanne, Switzerland, Oct. 26–29, 1997, vol. 2, pp. 518–521.

[18] K. Sobottka and I. Pitas, "Face localization and facial feature extraction based on shape and color information," in *Proc. IEEE Int. Conf. Image Process.*, Lausanne, Switzerland, Sep. 16–19, 1996, vol. 3, pp. 483–486.

[19] C. Kotropoulos and I. Pitas, "Rule-based face detection in frontal views," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Munich, Germany, Apr. 21–24, 1997, vol. 4, pp. 2537–2540.

[20] L. Yin and A. Basu, "Nose shape estimation and tracking for model-based coding," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Salt Lake City, Utah, May 7–11, 2001, vol. 3, pp. 1477–1480.

[21] L. G. Farkas, *Anthropometry of the Head and Face in Medicine*. New York: Elsevier, 1981.

[22] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[23] (2007). FLIR Systems, 25 Esquire Road, North Billerica, MA 01862 [Online]. Available: http://www.flir.com

[24] F. Martini, W. Ober, C. Garrison, K. Welch, and R. Hutchings, *Fundamentals of Anatomy and Physiology*, 5th ed. Upper Saddle River, NJ: Prentice-Hall, 2001, ch. 23.

[25] I. Pavlidis and J. Levine, "Monitoring of periorbital blood flow rate through thermal image analysis and its application to polygraph testing," in *Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Istanbul, Turkey, Oct. 25–28, 2001, vol. 3, pp. 2826–2829.

Jin Fei received the B.E. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1996, the M.S. degree in computer science and engineering from the State University of New York at Buffalo, Buffalo, and the Ph.D. degree from the Computer Science Department, University of Houston, Houston, TX, in December 2007.

He is currently with Halliburton, Inc., Houston. His research interests include biomedicine, computer vision, and image/signal processing. His research was supported by the National Science Foundation and the National Institutes of Health.

Dr. Fei was a recipient of the 2006 Scholarship and the Best Poster Awards from the Department of Computer Science, University of Houston.

Ioannis Pavlidis (S'85–M'87–SM'00) received the B.Sc. degree in electrical engineering from Democritus University of Thrace, Orestiada, Greece, in 1987, the M.Sc. degree in robotics from the University of London, London, U.K., in 1989, and the M.Sc. and Ph.D. degrees in computer science from the University of Minnesota, Minneapolis, in 1995 and 1996, respectively.

He is currently the Eckhard Pfeiffer Professor of Computer Science and the Director of the Computational Physiology Laboratory, University of Houston, Houston, TX. His research was supported by multiple federal agencies, including the National Science Foundation, the Department of Defense, the Department of Human Services, the National Institutes of Health, and corporate sources. He is the author or coauthor of many journal articles and books on the topics of computational physiology, computational psychology, computer vision, and pattern recognition. He is well known for his work on stress quantification, which appeared in *Nature* and *Lancet*, and received world-wide scientific and media attention.

Dr. Pavlidis is a Fulbright Fellow of the IEEE. He has also established several well-known IEEE Conferences and Workshops, including the IEEE Advanced Video and Signal Based Surveillance Conference.